DOCUMENT RESUME

ED 386 021 FL 023 058

AUTHOR	Kubota, Mikio
TITLE	The Garden Path Technique: Is It Really Effective?
PUB DATE	95
NOTE	30p.
PUB TYPE	Journal Articles (080)
JOURNAL CIT	Working Papers of Chofu Gakuen Women's Junior College; v27 p21-48 1995
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Classroom Research; Comparative Analysis; *English (Second Language); *Error Correction; *Grammar; *Grammatical Acceptability; Higher Education; Instructional Effectiveness; Language Research; Language Tests; Prevention; Second Language Instruction; *Second Languages; Teaching Methods; Testing; Two Year College Students
IDENTIFIERS	*Garden Path Technique; Japanese People; Passive Voice

ABSTRACT

This study replicating two earlier studies investigated the relative effectiveness of two techniques for teaching English grammar. In the "garden path" method, learners are induced to produce errors, which are then corrected. The error avoidance method prevents error production by means of grammatical explanations. Subjects were 56 Japanese junior college students of English as a Foreign Language. Both the experimental group (garden path treatment) and control group (error avoidance treatment) were administered two kinds of tests, one in which they were to make grammaticality judgments of the if-clause, and one in which they rewrote sentences using the passive voice correctly. The two tests gave mixed results: in the first, no statistical difference emerged between the two groups; in the second, the experimental group performed better. It also appeared that both treatments had a temporary effect in the first test, while both had long-term effects in the other test. It is concluded that both techniques may work well, at least in the short term, for this population. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 386 021

The Garden Path technique: is it really effective?

Mikio Kubota

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Mikio Kubota

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Working Papers of Chofu Gakuen Women's Junior College
1995, Vol 27, pp 21-48.
(March 20, 1995)
Kanagawa, Japan

FL023058

BEST COPY AVAILABLE

The Garden Path technique : is it really effective ?

Mikio Kubota

ABSTRACT

This research replicates two studies by Tomasello and Herron (1988, 1989) in order to investigate whether the garden path technique (inducing learners to produce errors and then correcting them) would outperform the error avoidance treatment (preventing errors by means of grammatical explanations). Fifty-six Japanese junior college students of EFL participated in this study. Both the experimental groups (the garden path treatment) and the control groups (the error avoidance treatment) were given the two kinds of tests in each test session (Pre-test session, Post-test 1 session, and Post-test 2 session) : Test (A), which required the subjects to make grammatical judgment of the *if* clause, and Test (B), which asked them to rewrite sentences using the passive voice correctly.

The findings resulting from this classroom research indicate :

(1) The two tests gave mixed results : in Test (A) there was no statistically significant difference found between the garden path treatment and the error avoidance treatment, whereas in Test (B) the garden path treatment was more effective than the error avoidance treatment.

(2) Both treatments had temporary effects of instruction in Test (A), while both treatments had long-term effects of instruction in Test (B).

These results pedagogically imply that teachers should keep in mind that both the garden path technique and the error avoidance treatment may work well at least in the short term, as far as Japanese college students are concerned.

1. INTRODUCTION

Teachers must make appropriate methodological decisions in every stage of a lesson. Therefore, the question "what kind of formal instruction works best?" concerns language teachers. In particular, they want to know whether error correction may facilitate learners' acquisition of correct forms.

An important distinction was made between a focus on *forms* and a focus on *form* (Long 1988, 1991). A focus on forms refers to "instruction that seeks to isolate linguistic forms in order to teach and test them one at a time" (Ellis 1994 : 639). This approach occurs when students "receive regular grammar lessons directed at specific items" (Ellis et al. 1994 : 10). On the contrary, a focus on form is a feature that involves "alternating in some principled way between a focus on meaning and a focus on form" (Long 1991 : 47). Learners "receive negative feedback on specific linguistic features in the course of trying to communicate" (Ellis et al. 1994 : 10). Long (1991 : 47) argues that instruction with a focus on forms is counter-productive and instruction with a focus on form produces a faster rate of learning and higher levels of ultimate attainment than instruction with no focus on form.

Recently, the relative efficacy of different types of formal instruction has been investigated. Doughty (1991) looked at the acquisition of relative clauses by 20 intermediate-level ESL learners, comparing mean-oriented instruction and rule-oriented instruction. The results revealed that both the mean-oriented instructional group and the rule-oriented instructional group outperformed the control group with respect to gains in relativization, and that the mean-oriented group better facilitated comprehension of the reading texts than was the case in either the rule-oriented or control groups. An apparent advantage for the mean-oriented group, that is, the *focus-on-form* group, with respect to comprehension was "attributed to the apparently successful combination of a focus on meaning and the bringing to prominence of the linguistic properties of relativization in the mean-oriented treatment" (Doughty 1991 : 463).

Tomasello and Herron (1988, 1989) explored the question of when attention to form may be most useful to learners. They compared two methods for teaching grammatical exceptions and translation to beginning

level English-speaking learners of French as a foreign language. In the 1988 study, in one condition the exception was taught to the students as an exception, while in the other condition, which Tomasello and Herron referred to as "the Garden Path" treatment, students were led to make overgeneralization errors and then it was corrected on the spot by the teacher. In the 1989 study, in one condition students were simply given the French form and were told that it differed from the English pattern, whereas in the Garden Path treatment transfer errors were induced and corrected on the spot by the teacher. The results of these studies revealed that the garden path treatment was more effective. Tomasello and Herron provided cognitive and motivational explanations for the results: learners were presented with the opportunity to make a cognitive comparison between their own system and the correct utterance (Tomasello and Herron 1989: 394). Moreover, "the disconfirmed expectation experienced by learners in cases of error correction heightens their motivation and/or curiosity to discover the differentiating features of the rule and its exception" (Tomasello and Herron 1988: 244).

Beck and Eubank (1991) criticized Tomasello and Herron (1989) mainly on the following methodological grounds:

1. the problems of the selection procedures for the subjects and the structures.
2. only one item tested for each structure on each test.
3. the heterogeneity of the structures investigated.
4. the lack of a pre-test.
5. the timing differentials among the three tests in each of the two classes.
6. only one task used to elicit data that will not generalize to other types of performance.
7. a question of researcher expectancy, when the researcher is personally involved in both the control and the experimental conditions.

Ellis et al. (1994: 13) add another critique:

8. no information provided on individual's performance.

Tomasello and Herron (1991) rebutted as follows:

1. An experimental investigation was performed within the context of

a naturally functioning classroom, which required a number of compromises in experimental design.

2. Technically, the results may be generalized to this population of subjects and structures only, that is, a pedagogically based selection criterion.
3. Student differences in previous experience with French are not a confound, but merely "noise" in the data.
4. Differences in the length of time between teaching and testing for different structures are taken care of by the counterbalanced design.

This study will address one of the important issues in classroom research - the role of negative feedback. It is a replication study of two studies by Tomasello and Herron (1988, 1989) for the purpose of investigating whether the garden path technique may be really beneficial to learners in EFL situations. There has been only one replication study (Ellis et al. 1994) on this topic, to the best of the present researcher's knowledge.

2. THE STUDY

The main purpose of this research is to examine whether the garden path technique may be more efficient than a traditional Error Avoidance method of instruction. The research questions are as follows:

1. What kinds of instructions will lead to the reformulations of subjects' grammatical knowledge?
2. What kinds of instructions will help subjects retain their grammatical knowledge?

2.1 HYPOTHESES

Hypotheses 0 and 1 are concerned with Research Question "1", while Hypotheses 2-5 are related to Research Question "2".

H0 : There would be no statistically significant difference in accuracy of responses between the experimental groups and the control groups.

The test scores would result in no difference between the experimental groups and the control groups. If this null hypothesis proves incorrect, the alternative

hypothesis is stated as follows:

H1: The experimental groups would outperform the control groups in the Post test 1 (immediate post test) and the Post test 2 (dalayed post test).

It is predicted, based on Tomasello and Herron (1988, 1989), that the garden path technique as negative feedback would be beneficial in inducing a positive learning effect. Tomasello and Herron (1988, 1989) found that students learned the grammatical exceptions and translation better in the experimental group (the garden path condition) than the control group, and that this advantage persisted thoughtout the semester-long course. It should be noted, however, that the same result was not obtained in the replication study carried out by Ellis et al. (1994). In spite of these mixed results obtained previously, the hypothesis that the garden path technique would be more effective than the error avoidance approach will be formed in order to investigate the efficacy of the garden path technique.

H2: There would be no statistically significant difference in accuracy of responses among the Pre test, the Post test 1, and the Post test 2.

This is also a null hypothesis, since no theory or previous research explains the differences among the Pre test, the Post test 1, and the Post test 2. The following alternative hypotheses 3-5 about the relationship between instruction and L2 learning will be tested in case the null hypothesis is rejected.

H3: Both experimental groups and control groups would respond significantly better in the Post test 1 than in the Pre test.

H4: Both experimental groups and control groups would respond significantly better in the Post test 2 than in the Pre test.

It is predicted that there would exist a positive effect of instruction on learning in both the experimental groups and the control groups. Therefore, the subjects receiving the garden path technique and the subjects receiving the error avoidance treatment would respectively get higher scores in the Post tests 1 and 2 than in the Pre test.

H5: Both experimental groups and control groups would respond significantly better in the Post test 1 than in the Post test 2.

Concerning the durability of instructional effects, Ellis (1993 : 17, 1994 : 637-638) states that the following three points should be taken into consideration in determining the durability:

(a) whether students have opportunities for input and output of the target structures, after the instruction is over,

(b) what a nature of the linguistic feature itself is (e.g., developmental (e.g., questions) or variational (e.g., adverbs) (see Meisel et al. 1981)),

(c) how students perceive the importance of a given grammatical feature.

In the present research, there was no class between the treatment and the delayed post test (Post test 2), and the subjects were requested not to study the target structures outside class. Hence, it seems that they had no opportunities to hear and use the structures. It is hypothesized that the effect of instruction would begin to disappear.

2.2 PROCEDURES

2.2.1 Subjects

A total of 56 Japanese women's junior college students were the subjects of the current research. They had studied EFL in instructional settings for six or seven years. They had already studied the basic usages of the structures investigated in this study, while they were in senior high school.

Intact classes were used for this research, since it was impossible to assign students randomly to the experimental and control groups. An additional 30 students had to be excluded in the analysis, because they missed one or more of the following : the Pre test, the treatment, the Post test 1, and the Post test 2.

2.2.2 Test Items

There were two grammatical items investigated : *whether / if* clauses and passive voices.

[Test (A)]

In Test (A), *whether / if* clauses were the target structure. The conjunctions *whether* and *if* are not always interchangeable (see Quirk et al. 1985: 1053-1054, Declerck 1991: 527-528, for more information). Both *whether* and *if* can normally introduce an object clause, as in (a):

(a) I asked him *whether / if* he would join us.

However, the *if*-clause cannot introduce a subject clause, a subject complement clause, and the complement of a preposition, as in (b), (c), and (d), respectively:*

(b) **If* John went to the park was not clear to me.

(c) *My main problem was *if* I could save $1000.

(d) *I didn't answer the question of *if* he had made a mistake.

The subjects were required to judge whether a sentence is grammatically correct. 10 items were given to the subjects, but the results of 6 items in each test session (Nos. 2, 3, 5, 7, 8, and 10 in Pre test; Nos. 2, 3, 5, 7, 9, and 10 in Post test) were examined for the purpose of this study.

[Test (B)]

In Test (B), the usage of the passive voice was examined. One type of grammatical exception was included in Test (B): if a causative verb or a perception verb with a bare infinitive clause is passivized, it requires a '*to*' infinitive clause, as in (e):

(e) One gentleman was seen *to* enter the bank after 8 o'clock.

The subjects were asked to rewrite a sentence using a passive voice. 15 items were given to the subjects, but the results of 5 items in each test session (Nos. 4, 7, 9, 12, and 14 in Pre test; Nos. 3, 7, 10, 13, and 15 in Post test) were examined for the purpose of the study.

In this experiment, there were three test sessions: Pre-test session, Post-test 1 session, and Post test 2 session. Both Test (A) and Test (B) were administered in each session. The same test items were used during three session, and Post-test 2 was the same test as Post-test 1. The test items used in the Pre-test were merely rearranged in Post-tests.

2.2.3 Research Design

Taking into account the above-mentioned eight criticisms of Tomasello and Herron (1989) raised by Beck and Eubank (1991) and Ellis et al. (1994), the following research design was created in order to ensure methodological validity, based on Ellis et al. (1994: 13):

1. there was a pre-test.
2. two kinds of tests were given to each subject.
3. the target structures were tested more thoroughly in each test.
4. there was no variability in the amount of elapsed time between treatment and testing.
5. results for individual learners as well as groups were reported and examined.

[Session 1: Pre-test session]

All the subjects in the four classes were given the 15-minute Pre-test (see Appendix 2).

[Session 2: Treatment session]

Four classes were randomly assigned to two garden path technique (experimental) groups and two Error Avoidance (control) groups. The two experimental groups consisted of 28 students, and the two control groups consisted of 28 students. Only one Japanese-speaking teacher of EFL (the present researcher) provided negative feedback in this experiment in order to control variables (feedback time on task, feedback contents), and feedback was given to each group collectively in Japanese, 20 minutes after the Pre-test during the same class period.

Test (A):

In both groups, grammatical explanations were first provided in Japanese as follows: the teacher wrote up the target structures used in the tests, pointing out that it is grammatical to use *whether* or *if* in the following sentences.

(1) a. I wonder *whether* you like beer.

b. I wonder *if* you like beer.

(2) a. He asked *whether* you could attend the meeting.

b. He asked *if* you could attend the meeting.

In the experimental group, the teacher asked the subjects to paraphrase the following two following sentences using "*if*" where it is possible :

(3) a. *Whether* he should go alone is still undecided.

(4) a. The problem is *whether* she should go to the party alone.

The data were collected so as to analyze whether the subjects actually made the expected errors or not. After the teacher circulated in class, the ungrammatical sentences most subjects produced were written up on the blackboard and were corrected on the spot by the teacher.

In the control group, two more examples were added on the board :

(3) a. *Whether* he should go alone is still undecided.

b. **If* he should go alone is still undecided.

(4) a. The problem is *whether* she should go to the party alone.

b. *The problem is *if* she should go to the party alone.

In both groups, the teacher explained in Japanese that "*if*" can introduce only an object clause such as (1) and (2), so that (3b) as a subject clause and (4b) as a subject complement are ungrammatical. Thus, the two groups differed in the following one respect : whether the subjects were provided the opportunity to commit the overgeneralization errors.

After the treatments, the Post test 1 of Test (A) was given.

Test (B):

After the Post test 1 of Test (A) was finished, the treatments in Test (B) were given. The treatments started in this same way in both the garden path groups and error avoidance groups. The teacher wrote the two target sentences of the active voice and the passive voice, and explained how to make the passive voice from the active one as follows :

(5) a. She asked me to carry the basket.

b. I was asked to carry the basket.

(6) a. I found him working at the desk.

b. He was found working at the desk.

In the experimental groups, the teacher required the subjects to paraphrase the following two sentences of the passive voice:

(7) a. John made her write the letter.

(8) a. We heard something fall onto the floor.

The subjects were expected to make overgeneralization errors. The data were again gathered in order to analyze the responses. After circulating in class, the ungrammatical sentences which some subjects produced were written up on the blackboard and were corrected on the spot by the teacher.

In the control groups, the following two more examples were added on the board for explanations of grammatical exceptions.

(7) a. John made her write the letter.

b. She was made *to* write the letter.

(8) a. We heard something fall onto the floor.

b. Something was heard *to* fall onto the floor.

In both groups, the teacher cautioned in Japanese that if a causative verb or a perception verb with a bare infinitive clause is passivized, it requires a "*to*" infinitive clause, as in (7b) and (8b). Thus, the two groups differed in the following one respect: whether the subjects were provided the opportunity to commit the overgeneralization errors.

After the treatments, the Post-test 1 of Test (B) was given.

[Session 3: Post-test 1 session]

Test (A) and Test (B) were given separately during the same class period as the Pre-test (see Appendix 2): Test (A) took 5 minutes, and Test (B) was 10 minutes long.

[Session 4: Post-test 2 session]

After one week passed, all the subjects were given the Post-test 2 (see Appendix 2), which was the same test as Post-test 1.

2.3 Data Analysis

A .05 level of significance ($\alpha = .05$) was selected. The *z* test and the two-way repeated measures ANOVA were employed to analyze the data.

3. RESULTS

3.1 Test (A)

The full mark of the structures examined in Test (A) was 6 points. The following Table 1 shows means and standard deviations by the experimental groups and the control groups. Means of correct responses in the Pre-test are displayed in Figure 1.

Table 2 indicates the results of the two-way repeated measures ANOVA for Test (A). The results show that only session differences were statistically

Table 1 : Means and standard deviations by group and session in Test (A)

Group	n	Mean	Standard Deviation
[Pre-test]			
Experimental	28	3.46	1.30
Control	28	3.57	1.74
[Post-test 1]			
Experimental	28	4.54	1.12
Control	28	4.29	1.22
[Post-test 2]			
Experimental	28	3.64	1.63
Control	28	4.03	1.40

Figure 1 : Means of correct responses in Test (A)

significant. It implies that only sessions influenced the results.

Multiple comparisons (using LSD: least square differences) were not made in order to determine which groups significantly differed from each other, since group differences were not significant. Hence, the null hypothesis (H0) was supported, so that Hypothesis 1 did not require testing. Moreover, it was confirmed that there was no difference between the experimental groups and the control groups in the Pre test, otherwise further analyses would not have been made; any comparative effects due to treatment were not related to prior knowledge or language ability of any one group.

Table 2. Results of two way repeated measures ANOVA for Test 'A'

Source	SS	df	MS	F	p
Between subjects	155.32	55			
Groups	0.29	1	0.29	0.10	ns
Subjects within groups	155.03	54	2.87		
Within subjects	172.67	112			
Session	22.90	2	11.45	8.42	.05
Groups by session	2.91	2	1.46	1.07	ns
Residual	146.86	108	1.36		
			$F_{2,54}$ .05	4.0848	
			$F_{2,108}$ .05	3.1504	

The results of between session comparisons of means for Test 'A' are shown in Table 3. Table 3 indicates which session were significantly different from each other. Both the experimental groups and the control groups performed better in the Post test 1 than in the Pre test, and only the experimental groups did better in the Post test 1 than in the Post test 2. Therefore, the null hypothesis (H2) was rejected, thereby Hypothesis 3 was supported, but Hypothesis 4 was not supported. Hypothesis 5 was partly supported, because only the experimental groups responded better in the Post test 1 than in the Post test 2.

3.2 Test (B)

The full mark of items examined in Test 'B' was 5 points. Table 4 shows

Table 3 Between session comparisons of means for Test 'A'

Mean	Session	Pre test	Post test 1	Post test 2
Experimental				
3.46	Pre test			
4.54	Post test 1	*		
3.64	Post test 2		*	
Control				
3.57	Pre test			
4.29	Post test 1	*		
4.03	Post test 2			

* p < .05

Table 4 Means and standard deviations by group and session in Test 'B'

Group	n	Mean	Standard Deviation
[Pre-test]			
Experimental	28	1.46	2.03
Control	28	1.11	1.70
[Post-test 1]			
Experimental	28	4.82	0.47
Control	28	4.14	1.30
[Post-test 2]			
Experimental	28	4.36	0.97
Control	28	3.79	1.63

the respective means and standard deviations by each group and session in Test 'B'. Means of correct responses in Test 'B' are illustrated in Figure 2. Table 5 demonstrates the results of the two way repeated measures ANOVA for Test 'B'. As indicated in Table 5, group differences and session differences were statistically significant. Furthermore, the group by session interaction was statistically significant. Since the main effects for groups and sessions and the interaction were significant, groups and sessions influenced the results dependently.

Figure 2 : Means of correct responses in Test (B)

Table 5 : Results of two way repeated measures ANOVA for Test (B)

Source	SS	df	MS	F	p
Between subjects	159.85	55			
Groups	12.05	1	12.05	4.40	<.05
Subjects within groups	147.80	54	2.74		
Within subjects	484.00	112			
Session	338.73	2	169.37	282.28	<.05
Groups by session	80.70	2	40.35	67.25	<.05
Residual	64.57	108	0.60		

$F_{(2,54)}(.05) = 4.0848$

$F_{(2,108)}(.05) = 3.1504$

Multiple comparisons were made in order to decide which groups significantly differed from each other. The results of between group comparisons are shown in Tables 6, 7, and 8. In the Pre test, no statistically significant difference was found between the experimental groups and the control groups, so that further analyses continued. It was confirmed that any comparative effects due to treatment were not related to prior knowledge or language ability of any one group.

In both the Post test 1 and the Post test 2, the experimental groups significantly outperformed the control groups. Therefore, the null hypothesis

Table 6 Between group comparisons of means for Pre test of Test (B)

Mean	Group	Experimental	Control
1.46	Experimental		
1.11	Control		

* p< .05

Table 7 Between group comparisons of means for Post test 1 of Test (B)

Mean	Group	Experimental	Control
4.82	Experimental		
4.14	Control	*	

* p< .05

Table 8 Between group comparisons of means for Post test 2 of Test (B)

Mean	Group	Experimental	Control
4.36	Experimental		
3.79	Control	*	

* p< .05

Table 9 Between session comparisons of means for Test (B)

Mean	Session	Pre test	Post test 1	Post test 2
Experimental				
1.46	Pre test			
4.82	Post test 1	*		
4.36	Post test 2	*	*	
Control				
1.11	Pre test			
4.14	Post test 1	*		
3.79	Post test 2	*		

* p< .05

(H0) was rejected, thereby Hypothesis 1 was supported.

The results of between session comparisons of means in Test (B) are shown in Table 9. Table 9 indicates which sessions were significantly differ-

BEST COPY AVAILABLE

ent from each other. Both the experimental groups and the control groups performed better in the Post tests 1 and 2 than in the Pre test. Only the experimental groups did better in the Post test 1 than in the Post test 2. Thus, the null hypothesis (H2) was rejected, thereby Hypotheses 3 and 4 were supported, but Hypothesis 5 was supported only in the experimental groups.

4. DISCUSSION

The following hypotheses were tested in this study:

H0: There would be no statistically significant difference in accuracy of responses between the experimental groups and the control groups.

H1: The experimental groups would outperform the control groups in the Post test 1 and the Post test 2.

H2: There would be no statistically significant difference in accuracy of responses among the Pre test, the Post test 1, and the Post test 2.

H3: Both experimental groups and control groups would respond significantly better in the Post test 1 than in the Pre test.

H4: Both experimental groups and control groups would respond significantly better in the Post test 2 than in the Pre test.

H5: Both experimental groups and control groups would respond significantly better in the Post test 1 than in the Post test 2.

H0–H1:

In Test (A), which required the subjects to make grammatical judgments about the *if* clause, the null hypothesis (H0) was supported in that no statistically significant difference was found between the experimental groups and the control groups. In contrast, in Test (B), which asked the subjects to rewrite sentences with a passive voice, H1 was supported in that the garden path treatment outperformed the error avoidance treatment. Thus, Test (A) and Test (B) gave contradictory results.

H2–H5:

In Test (A), the null hypothesis (H2) was rejected, thereby Hypothesis 3 was supported in that both the experimental groups and the control groups performed better in the Post test 1 than in the Pre test. However, Hypothesis 4 was not supported. Hypothesis 5 was partly supported in that only the

experimental groups did better in the Post-test 1 than in the Post-test 2.

In Test (B), the null hypothesis (H2) was rejected, so that Hypotheses 3 and 4 were supported in that both the experimental groups and the control groups responded better in the Post-tests 1 and 2 than in the Pre-test, but Hypothesis 5 was supported only in the experimental groups. Only the experimental groups did better in the Post-test 1 than in the Post-test 2.

Table 10 summarizes the studies investigating the effect of the garden path technique, including the present study.

There are five possibilities that should be taken into account in explaining these mixed results of the effectiveness of the garden path technique, as mentioned in Ellis et al. (1994 : 17-18) :

(a) subjects' experiences in classroom learning

The subjects investigated in Tomasello and Herron (1988, 1989) were complete beginners, so that those who had little experience in classroom learning may have found it difficult to benifit from grammatical explanations. On the contrary, it seems that Japanese subjects may have experienced sufficient explanations for many years during high school, thereby they may have been in a better position to benefit from formal instruction than those subjects with little experience. However, the two studies examining Japanese learners gave contradictory results. It may be possible that the degree of learning experiences in class would influence the result. At any rate, this topic remains at issue.

(b) whether subjects make a cognitive comparison between their own system and that of native speakers in a class where a single grammatical structure is taught thoroughly (as in Ellis et al. (1994) and the present study) or in a class where a number of different structures are targeted (as in Tomasello and Herron (1988, 1989))

In the former type of the design, the present study showed that the garden path treatment was beneficial in Test (B), but not in Test (A) ; in the latter type of the design Tomasello and Herron (1988, 1989) found the same result. Therefore, this topic also needs further investigation.

(c) when an advantage for the garden path treatment becomes apparent

It is possible to state that a certain type of treatment has a short-term

Table 10 Studies investigating the effect of the garden path technique

Study	Type of classroom	Subjects	Proficiency	Data	Results
Tomasello and Herron (1988)	French as a foreign language in USA	Adults at university	Beginning	Fill in the blank tasks	·Garden path treatment > error avoidance treatment > overgeneralization error
Tomasello and Herron (1989)	French as a foreign language in USA	Adults at university	Beginning	Translation tests	·Garden path treatment > error avoidance treatment > transfer error
Ellis et al. (1994)					
Study 1	EFL in Japan	Adults at a private language school	Beginning	Sentence completion tasks	·Garden path treatment > error avoidance treatment ·Both treatments were temporarily effective in avoiding overgeneralization errors.
Study 2	EFL in Japan	Adults at university	Beginning	Grammaticality judgment tasks	·Garden path treatment > error avoidance treatment ·Neither treatment was temporarily effective in avoiding overgeneralization errors.
Kubota present study					
Test A	EFL in Japan	Adults at junior college	Intermediate	Grammaticality judgment tasks	·Garden path treatment > error avoidance treatment ·Both treatments were temporarily effective in avoiding overgeneralization errors.
Test B	EFL in Japan	Adults at junior college	Intermediate	Sentence rewrite tasks	·Garden path treatment > error avoidance treatment ·Both treatments were effective in the long term in avoiding overgeneralization errors.

> : more effective
= : not different from

effect, or leads to long-term retention of the correct forms, or has a latency effect in which there exists no immediate effect but the effect appears some time after the instruction. In Tomasello and Herron (1988, 1989), where there was no pre-test, the garden path treatment produced the long-term effect throughout the semester. Ellis et al. (1994), which administered only the immediate post-test, had no immediate effect. In the current research, both the garden path treatment and the error avoidance treatment were temporarily beneficial in avoiding overgeneralization errors in grammaticality judgment tasks of Test (A), but both treatments had the long-term effect in sentence-rewrite tasks of Test (B). It may be the case that the type of the instructional treatment determines when an effect will emerge.

(d) the difficulty of tests

It may be that the difficulty of tests given will influence the results. This possibility was not examined, because it is beyond the scope of the study.

(e) the learner factors

The results may rely on learning styles of individual learners and their motivations. This topic is not within the scope of the study, either.

The other possibilities are added here:

(f) the language features investigated

It is not clear whether a treatment is effective with some types of linguistic structures (e.g., syntactic, lexical, morphological), with some types of linguistic features (e.g., developmental or variational), or with some types of L2 processes (e.g., overgeneralization, transfer). This possibility should be taken into consideration in future research.

(g) how the target structures are taught

The results also depend on whether the treatment is provided "in the context of a naturally functioning classroom" (Tomasello and Herron 1991: 514) or in the context of a grammar-oriented instruction. Tomasello and Herron (1988, 1989) followed the former type, while Ellis et al. (1994) and the current research were based on the latter type. How the target structures are taught may be one of the factors to determine the efficacy of an instructional treament. This topic also remains at issue.

(h) how data are collected

Tomasello and Herron (1988, 1989) used fill in the blanks tests and translation tests, rejectively. In contrast, Ellis et al. (1994) employed sentence completion tests and grammaticality judgment tests, and the current research used grammaticality judgment tests and sentence rewrite tests. Futher research should include oral and written data to elicit productive and receptive knowledge of subjects.

The garden path treatment caused the subjects to construct incorrect hypotheses, but it does not follow that all the subjects actually do so during the treatment session, as Ellis (1994: 18) pointed out. In order to check this possibility, the post-hoc analysis was made. It reveals that the total number of subjects who were led to erroneous sentences owing to the garden path treatment was 50 in the two items of Test (A) and 16 in the two items of Test (B). This variability of numbers may rely on the task types and the difficulty of the target structures investigated. Test (A) seems easier because of the grammaticality judgments than Test (B), which asked the learners to rewrite sentences.

In contrast, the error avoidance treatment seeks to prevent learners from making errors, but it does not follow that at some stage of the instruction the learners did not construct a mental representation of the erroneous sentences, as Ellis et al. (1994: 18) stated. The data were not collected in the error avoidance treatment. Therefore, further research should investigate what goes in mind of the individual learners when confronted with a particular instructional experience (Ellis et al. 1994: 18) by conducting interviews with every learner or by making self-reports. How a mental representation of the erroneous sentences will influence the actual performance will be an interesting issue.

It may be important to point out the fact that both treatments had immediate effects in Test (A) and long-term effects in Test (B), resulting in beneficial instructions. It may be plausible that the grammaticality judgment test, employed in Test (A), gave variance because there was a 50% chance of getting correct responses. Thus, test methodologies may determine the durability of a treatment, besides the three points mentioned earlier by Ellis (1993:

17, 1994 : 637-638), such as (a) opportunities for input and output of the target structures, (b) the nature of the linguistic feature investigated, and (c) students' perception of the importance of the structure.

The results for individual learners demonstrate that in the experimental treatment of Test (A), there were two subjects (Nos. 7 and 11) whose scores on the grammaticality judgment test dropped considerably in the Post test 2 (see Appendix 1, Test (A)) ; this may be that they wrote down all the answers in the wrong section of the answer sheet.

5. CONCLUSION

The following major findings and pedagogical implications are presented, based on this classroom research.

(1) The two tests gave mixed results : in Test (A) there was no statistically significant difference found between the garden path treatment and the error avoidance treatment, whereas in Test (B) the garden path treatment was more effective than the error avoidance treatment.

(2) Both treatments had temporary effects of instruction in Test (A), while both treatments had long-term effects of instruction in Test (B).

Therefore, the answer to "Is the garden path technique effective?" is "Sometimes Yes, sometimes No" regarding an advantage for the garden path technique over the error avoidance treatment, and "Yes" concerning temporary effects of instruction. Teachers should keep in mind that inducing learners to produce an error which is then immediately corrected and preventing errors by means of grammatical explanations work well, at least in the short term, as far as Japanese college students are concerned.

The limitations of this study are as follows : (a) a small number of subjects (n=56), (b) the reseacher bias, (c) no oral data collected, and (d) no follow-up post-test that may be given one month or a few months after the treatment in order to check the long-term effect in a real sense.

The following various conditions should be taken into account in determining whether the garden path technique is beneficial : (a) subjects' experiences in classroom learning, (b) whether subjects make a cognitive comparison between their own system and that of native speakers in a class where a

single grammatical structure is taught thoroughly or in a class where a number of different structures are targeted, (3) when an advantage for the garden path treatment becomes apparent, (4) the difficulty of tests, (5) the learner factors, (6) the language features investigated, (7) how the target structures are taught, and (8) how data are collected.

Further research will include the research design where the experimental groups contain Group (A) : only negative feedback, Group (B) : negative feedback and grammatical explanations, and Group (C) : only grammatical explanations, and the control group has no treatment, in order to examine which of these types of information learners utilize, or in what way (see Ellis et al. 1994 : 19).

NOTE

1. Several researchers criticized the grammaticality judgment task (Akamatsu 1992 : 122) : (1) there may be a possibility that a subject would judge all sentences either grammatical or ungrammatical (White 1989), (2) there may be a possibility that a subject would judge sentences on the basis of different criteria (e. g., semantic or pragmatic criteria) from what an experimenter intends (White 1989, Schachter and Yip 1990), and (3) it is not clear whether a grammaticality judgment made by the subjects is based on their linguistic knowledge or on the processing difficulties (Schachter 1989, Schachter and Yip 1990). Ellis et al. (1994 : 21), too, argue that the reliability of the grammaticality judgment tests is doubtful, as it is difficult to produce consistent judgments.

REFERENCES

Akamatsu, T. 1992. Methodological considerations of UG based studies in second language acquisition : a brief review. *CASELE Research Bulletin*, 22, 117-125. Chugoku Academic Society of English Language Education.

Beck, M. and Eubank, L. 1991. Acquisition theory and experimental design. *Studies in Second Language Acquisition*, 13, 73-76.

Beebe, L. M. (ed.) 1988. *Issues in second language acquisition : multiple perspectives*. New York : Newbury House.

de Bot, K., Ginsberg, R. B., and Kramsch, C. (eds.) 1991. *Foreign language research in cross-cultural perspective*. Amsterdam : John Benjamins.

Declerck, R. 1991. *A comprehensive descriptive grammar of English*. Tokyo : Kaitakusha.

Doughty, C. 1991. Second language instruction does make a difference : evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, 13, 4, 431-439.

Ellis, R. 1993. Grammar teaching and grammar learning. *Temple University Japan Research Studies in TESOL*, 1, 1, 3-27.

Ellis, R. 1994. *The study of second language acquisition*. Oxford : Oxford University Press.

Ellis, R., Rosszell, H. R., and Takashima, H. 1994. Down the Garden Path : another look at negative feedback. *JALT Journal*, 16, 1, 9-24.

Long, M. H. 1988. Instructed interlanguage development. In Beebe, L. M. (ed.), 115-141.

Long, M. H. 1991. Focus on form : a design feature in language teaching methodology. In de Bot, K. et al., 39-52.

Meisel, J., Clahsen, H., and Pienemann, M. 1981. On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition*, 3, 1, 109-135.

Quirk, R., Greenbaum, S., Leech, G., and Startvik, J. 1985. *A comprehensive grammar of the English language*. London : Longman.

Schachter, J. 1989. A new look at an old classic. *Second Language Research*, 5, 1, 30-47.

Schachter, J. and Yip, V. 1990. Grammaticality judgements : why does anyone object to subject extraction ? *Studies in Second Language Acquisition*, 12, 379-392.

Tomasello, M. and Herron, C. 1988. Down the Garden Path : inducing and correcting overgeneralization errors in the foreign language classroom. *Applied Psycholinguistics*, 9, 3, 237-246.

Tomasello, M. and Herron, C. 1989. Feedback for language transfer errors : the Garden Path Technique. *Studies in Second Language Acquisition*, 11, 4, 385-395.

Tomasello, M. and Herron, C. 1991. Experiments in the real world : a reply to Beck and Eubank. *Studies in Second Language Acquisition*, 13, 4, 513-517.

White, L. 1989. *Universal grammar and second language acquisition*. Amsterdam : John Benjamins.

BEST COPY AVAILABLE

APPENDIX 1

Results for individual subjects in Test (A) and Test (B)

Test (A)

Experimental Groups (n = 28)				Control Groups (n = 28)			
Subject	Pre test	Post test 1	Post test 2	Subject	Pre test	Post test 1	Post test 2
No. 1	4	6	3	No. 1	4	5	6
No. 2	2	5	3	No. 2	3	4	5
No. 3	4	6	4	No. 3	5	5	4
No. 4	5	4	7	No. 4	6	6	6
No. 5	6	5	5	No. 5	4	5	4
No. 6	4	4	4	No. 6	5	5	4
No. 7	4	6	6	No. 7	4	3	3
No. 8	3	4	5	No. 8	4	2	3
No. 9	3	5	4	No. 9	4	5	2
No.10	4	5	5	No.10	5	4	5
No.11	4	5	6	No.11	4	4	5
No.12	2	4	4	No.12	2	3	2
No.13	3	2	2	No.13	3	3	3
No.14	2	4	6	No.14	3	6	5
No.15	4	5	3	No.15	5	4	5
No.16	3	2	3	No.16	3	6	3
No.17	5	4	4	No.17	3	3	3
No.18	3	4	3	No.18	3	4	5
No.19	6	5	4	No.19	3	4	2
No.20	3	6	6	No.20	1	1	1
No.21	4	5	6	No.21	4	4	5
No.22	7	4	3	No.22	4	6	6
No.23	2	3	3	No.23	6	4	4
No.24	4	6	6	No.24	2	4	2
No.25	5	6	4	No.25	4	6	5
No.26	4	5	4	No.26	1	5	3
No.27	4	4	6	No.27	1	3	6
No.28	5	5	4	No.28	4	4	4

Test (B)

Experimental Groups (n = 28)				Control Groups (n = 28)			
Subject	Pre test	Post test 1	Post test 2	Subject	Pre test	Post test 1	Post test 2
No. 1	1	4	5	No. 1	0	5	5
No. 2	1	5	4	No. 2	0	5	5
No. 3	2	5	5	No. 3	1	2	3
No. 4	1	4	4	No. 4	3	5	5
No. 5	0	5	5	No. 5	0	5	5
No. 6	1	5	4	No. 6	5	5	5
No. 7	0	4	4	No. 7	0	0	0
No. 8	2	5	5	No. 8	0	5	4
No. 9	0	5	4	No. 9	2	5	5
No. 10	2	5	4	No. 10	0	5	5
No. 11	1	5	4	No. 11	0	5	0
No. 12	2	4	4	No. 12	1	5	5
No. 13	0	5	5	No. 13	0	4	4
No. 14	2	5	5	No. 14	0	5	2
No. 15	0	5	5	No. 15	0	4	5
No. 16	0	5	5	No. 16	0	2	5
No. 17	0	5	4	No. 17	0	5	2
No. 18	0	5	5	No. 18	1	5	5
No. 19	0	5	5	No. 19	2	5	5
No. 20	2	5	5	No. 20	0	5	5
No. 21	2	5	5	No. 21	0	4	5
No. 22	5	5	5	No. 22	2	5	5
No. 23	4	5	4	No. 23	5	4	5
No. 24	2	5	5	No. 24	4	5	5
No. 25	3	5	5	No. 25	5	4	4
No. 26	0	5	5	No. 26	0	2	0
No. 27	0	5	5	No. 27	0	5	5
No. 28	0	5	5	No. 28	0	5	4

APPENDIX 2

[Pre-test]

Test (A) Put an ○ if the following sentence is grammatically correct, and put an X if it isn't.

1. (a) I wonder whether you like beer.
   (b) I wonder if you like beer.
2. (a) The result depends on whether we will work hard.
   (b) The result depends on if we will work hard.
3. (a) Whether he should go alone is still undecided.
   (b) If he should go alone is still undecided.
4. (a) I don't know whether he'll be staying here.
   (b) I don't know if he'll be staying here.
5. (a) The question is whether the typhoon will come to our town.
   (b) The question is if the typhoon will come to our town.
6. (a) Tell me whether you are warm enough.
   (b) Tell me if you are warm enough.
7. (a) Whether the man agrees is not clear.
   (b) If the man agrees is not clear.
8. (a) The problem is whether she should go to the party alone.
   (b) The problem is if she should go to the party alone.
9. (a) He asked whether you could attend the meeting.
   (b) He asked if you could attend the meeting.
10. (a) There is some doubt as to whether he is guilty.
    (b) There is some doubt as to if he is guilty.

Test (B) Rewrite the following sentences with the passive voice.

1. She asked me to carry the basket.
2. I found him working at the desk.
3. He saw the house destroyed.
4. John made her write the letter.
5. They believed Ken to be innocent.

6. We told him to play the piano.
7. She saw Mike break the vase.
8. He heard his teacher singing.
9. I watched the boy kick the ball.
10. I assumed him to be honest.
11. They requested me to take the lead.
12. We heard something fall onto the floor.
13. My father allowed me to drink wine.
14. The woman made him sign a contract.
15. We thought Bob to be a nice man.

[Post-test]

NB. Post-test 2 was the same as Post-test 1.

Test (A) Put an ○ if the following sentence is grammatically correct, and put an X if it isn't.

1. a. I don't know whether he'll be staying here.
   b. I don't know if he'll be staying here.
2. a. Whether the man agrees is not clear.
   b. If the man agrees is not clear.
3. a. There is some doubt as to whether he is guilty.
   b. There is some doubt as to if he is guilty.
4. a. He asked whether you could attend the meeting.
   b. He asked if you could attend the meeting.
5. a. The result depends on whether we will work hard.
   b. The result depends on if we will work hard.
6. a. Tell me whether you are warm enough.
   b. Tell me if you are warm enough.
7. a. The question is whether the typhoon will come to our town.
   b. The question is if the typhoon will come to our town.
8. a. I wonder whether you like beer.
   b. I wonder if you like beer.
9. a. The problem is whether she should go to the party alone.
   b. The problem is if she should go to the party alone.
10. a. Whether he should go alone is still undecided.
    b. If he should go alone is still undecided.

Test (B) Rewrite the following sentences with the passive voice.

1. We told him to play the piano.
2. He heard his teacher singing.
3. John made her write the letter.
4. I assumed him to be honest.
5. She asked me to carry the basket.
6. He saw the house destroyed.
7. I watched the boy kick the ball.
8. I found him working at the desk.
9. They believed Ken to be innocent.
10. She saw Mike break the vase.
11. We thought Bob to be a nice man.
12. My father allowed me to drink wine.
13. The woman made him sign the contract.
14. They requested me to take the lead.
15. We heard something fall onto the floor.